

IFW

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/527, 552
Source: PCT
Date Processed by STIC: 12/08/2005

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 12/08/2005

PATENT APPLICATION: US/10/527,552

TIME: 10:41:00

Input Set : A:\32677_sequence listing.txt

Output Set: N:\CRF4\12082005\J527552.raw

3 <110> APPLICANT: Sandoz GmbH
 5 <120> TITLE OF INVENTION: Process for production of cephalosporin C
 7 <130> FILE REFERENCE: IB/G-32677A/BCK
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/527,552
 C--> 10 <141> CURRENT FILING DATE: 2005-03-11
 12 <160> NUMBER OF SEQ ID NOS: 21
 14 <170> SOFTWARE: PatentIn version 3.2
 16 <210> SEQ ID NO: 1
 17 <211> LENGTH: 526
 18 <212> TYPE: PRT
 19 <213> ORGANISM: Acremonium chrysogenum
 21 <400> SEQUENCE: 1
 23 Met Ala Ser Pro Ile Ala Ser Ala Ala Leu Lys Ala Arg Ile Arg Arg
 24 1 5 10 15
 27 Pro Ser Met Leu Lys Lys Leu Cys Lys Pro Gln Asp Leu Met His His
 28 20 25 30
 31 Phe Pro Asn Gly Ser Tyr Ile Gly Trp Ser Gly Phe Thr Gly Val Gly
 32 35 40 45
 35 Tyr Pro Lys Lys Met Pro Thr Tyr Met Ala Asp His Val Glu Gln Asn
 36 50 55 60
 39 Gly Leu Gln Gly Lys Leu Lys Tyr Ser Leu Phe Val Gly Ala Ser Ser
 40 65 70 75 80
 43 Gly Ala Glu Thr Glu Asn Arg Trp Ala Ser Leu Asp Met Ile Asp Arg
 44 85 90 95
 47 Arg Thr Pro His Gln Val Gly Lys Ala Ile Ser Lys Gly Ile Asn Glu
 48 100 105 110
 51 Gly Lys Ile His Phe Phe Asp Lys His Leu Ser Met Phe Pro Val Asp
 52 115 120 125
 55 Leu Val Tyr Gly Tyr Tyr Thr Lys Asp Arg Pro His Asn Lys Leu Asp
 56 130 135 140
 59 Val Val Val Val Glu Ala Thr Asp Ile Lys Glu Asp Gly Ser Ile Val
 60 145 150 155 160
 63 Pro Gly Ala Ser Val Gly Ala Thr Pro Glu Leu Ile Gln Met Ala Asp
 64 165 170 175
 67 Lys Ile Ile Ile Glu Val Asn Thr Ser Leu Pro Ser Phe Glu Gly Leu
 68 180 185 190
 71 His Asp Ile Thr Met Thr Asp Leu Pro Pro Leu Arg Lys Pro Tyr Leu
 72 195 200 205
 75 Val Met Gly Val Glu Asp Arg Ile Gly Arg Thr Ser Ile Pro Ile Asp
 76 210 215 220
 79 Pro Glu Lys Val Val Gly Ile Leu Glu Ser Asp Tyr Gln Asp Ala Thr
 80 225 230 235 240
 83 Ala Pro Asn Ala Glu Ala Asp Glu Ser Ala Asn Lys Ile Ala Gly His

(pg-6)

RAW SEQUENCE LISTING

DATE: 12/08/2005

PATENT APPLICATION: US/10/527,552

TIME: 10:41:00

Input Set : A:\32677_sequence listing.txt

Output Set: N:\CRF4\12082005\J527552.raw

```

84          245          250          255
87 Leu Ile Glu Phe Phe Glu His Glu Val Ala His Gly Arg Leu Pro Asn
88          260          265          270
91 Ser Leu Leu Pro Leu Gln Ser Gly Ile Gly Asn Val Ala Asn Ala Ile
92          275          280          285
95 Ile Gly Gly Leu Asp Asn Ser Asn Phe Arg Asn Leu Lys Val Trp Thr
96          290          295          300
99 Glu Val Ile Gln Asp Thr Phe Leu Asp Leu Phe Asp Ser Gly Arg Leu
100 305          310          315          320
103 Asp Phe Ala Thr Ala Thr Ser Ile Arg Phe Ser Pro Asp Gly Phe Arg
104          325          330          335
107 Arg Phe Tyr Asp Asn Trp Glu Ala Tyr Tyr Gly Lys Leu Leu Leu Arg
108          340          345          350
111 Ser Gln Gln Val Ser Asn Ser Pro Glu Ile Ile Arg Arg Leu Gly Val
112          355          360          365
115 Ile Ala Met Asn Thr Pro Val Glu Val Asp Ile Tyr Ala His Ala Asn
116          370          375          380
119 Ser Thr Cys Val Met Gly Ser Arg Met Leu Asn Gly Leu Gly Gly Ser
120 385          390          395          400
123 Ala Asp Phe Leu Arg Ser Ser Lys Tyr Ser Ile Met His Thr Pro Ser
124          405          410          415
127 Thr Arg Pro Ser Lys Thr Asp Pro His Gly Val Ser Cys Ile Val Pro
128          420          425          430
131 Met Cys Thr His Ile Asp Gln Thr Glu His Asp Leu Asp Val Ile Val
132          435          440          445
135 Thr Glu Gln Gly Leu Ala Asp Val Arg Gly Leu Ser Pro Arg Glu Arg
136          450          455          460
139 Ala Arg Val Ile Ile Lys Lys Cys Ala His Pro Val Tyr Gln Pro Ile
140 465          470          475          480
143 Leu Thr His Tyr Phe Glu Lys Ala Glu Ser Asp Cys Leu Arg Lys Gly
144          485          490          495
147 Trp Gly His Glu Pro His Leu Leu Phe Asn Ser Phe Asp Leu His Lys
148          500          505          510
151 Ala Leu Val Glu His Gly Ser Met Gln Lys Val Gly Gln Trp
152          515          520          525
155 <210> SEQ ID NO: 2
156 <211> LENGTH: 1765
157 <212> TYPE: DNA
158 <213> ORGANISM: Acremonium chrysogenum
160 <400> SEQUENCE: 2
161 atggcatcac caatagcctc tgccgccttc aaggcgcgca ttcgccgccc ctgatgctc 60
163 aagaagctgt gcaagcccca ggatttgatg catcacttcc ccaatggctc gtacattggc 120
165 tgggtccggt tcaccggcgt cggtacccg aagtgagttc caccgtcatc ccgctccaca 180
167 gtaggcgcag ccggcccgtt gacagtcccc gacaggaaaa tgccgacctc catggccgat 240
169 cacgtcgagc agaacggcct tcaggggcaag ctgaagtact cgctattcgt gggcgcatcg 300
171 tcgggtgctg agacagagaa tcgctgggcg tcgctcgaca tgattgatag gaggaccccg 360
173 catcaggctg gcaaggccat ctccaagggc atcaatgagg gcaagatcca cttcttcgac 420
175 aagcatctct ccattgtccc cgtggacctt gtatacgtag gtcaacgatg atcccttggc 480
177 atgtgcatgt actacgagta cctggcgcta acatccggtc agggctacta cacaaggat 540

```

RAW SEQUENCE LISTING

DATE: 12/08/2005

PATENT APPLICATION: US/10/527,552

TIME: 10:41:00

Input Set : A:\32677_sequence_listing.txt

Output Set: N:\CRF4\12082005\J527552.raw

```

179 agacccacaca acaagctgga cgtggtggtg gtggaggcca cgcacatcaa agaggacgga      600
181 agcattgtac ccggagcttc agtcggcgcg acccccgagc tcatccagat ggccgataag      660
183 gtgagcaatt tcgatttcta gcggaggggcg cagcaggacc tgacatctcc ctgtgcagat      720
185 cattatcgag gtcaaacacct cactgccttc attcgagggt ctccacgaca tcaccatgac      780
187 cgacctgccc ccgctacgga agccctatct cgtcatgggt gtcgaggacc gcacgcgcag      840
189 gacctctatc cctatcgacc ccgagaaggt tgtaggcatc ctcgaaatccg actaccagga      900
191 cgccactgcc cccaacgccg aggccgacga gagtgcgaac aagattgctg gccacttgat      960
193 tgagttcttc gagcacgagg tcgcccacgg ccgtctcccg aactccctcc ttcccctcca     1020
195 gtccggcatc ggcaacgctc ccaacgccat catcgggtggc ctcgacaact ccaacttccg     1080
197 caacctcaag gtctggactg aggttatcca ggacaccttc ctcgacctct tcgactcggg     1140
199 ccgcctcgac tttgccacgg ccacctctat ccgcttctcc cccgacggtt tccgcgggtt     1200
201 ctacgacaac tgggaggcct actacggcaa gctcctcctc cgcagccagc aggtgtccaa     1260
203 ctgcccagag atcatccgcc gccttgggtg cattgccatg aacacccccg tcgaggtcga     1320
205 catctacgcc cagccaact ccacctgcgt catgggctcg cgcagtctca acggcctggg     1380
207 cggctccgcc gacttccctg gctcctccaa gtactctatc atgcacaccc cgtccacccg     1440
209 cccctccaag accgacccgc accgcgctct gtgcatcggt cccatgtgca cccacatcga     1500
211 ccagactgag cagcacctcg acgtcatcgt caccgagcag ggcttgccg acgtgcgcgg     1560
213 cctgagcccc agggagaggg cccgcgtcat catcaagaag tgcgccacc cggctctacca     1620
215 gcccatcctg acccactact ttgagaaggc cgagagcgac tgcctacgca agggctgggg     1680
217 ccacgagccc catctgctct tcaactcggt tgacctgcac aaggccctcg tggagcacgg     1740
219 aagcatgcag aaggctcgggc agtgg                                     1765
222 <210> SEQ ID NO: 3
223 <211> LENGTH: 1578
224 <212> TYPE: DNA
225 <213> ORGANISM: Acremonium chrysogenum
227 <400> SEQUENCE: 3
228 atggcatcac caatagcctc tgccgccttc aaggcgcgca ttccgccccc ctcgatgetc      60
230 aagaagctgt gcaagcccca ggatttgatg catcacttcc ccaatggctc gtacattggc     120
232 tggtcgggt tcaccggcgt cggctacccg aagaaaatgc cgacctacat ggccgatcac     180
234 gtcgagcaga acggccttca gggcaagctg aagtactcgc tattcgtggg cgcacgtcg     240
236 ggtgctgaga cagagaatcg ctgggcgtcg ctcgacatga ttgataggag gaccccgcat     300
238 caggtcggca aggccatctc caagggcac aatgagggca agatccactt cttcgacaag     360
240 catctctcca tgttccccgt ggaccttgta tacggctact acacaaaagga tagaccacac     420
242 aacaagctgg acgtgggtgg ggtggaggcc accgacatca aagaggacgg aagcattgta     480
244 cccggagctt cagtcggcgc gacccccgag ctcatccaga tggccgataa gatcattatc     540
246 gaggtcaaca cctcactgcc ttcattcgag ggtctccacg acatcaccat gaccgacctg     600
248 cccccgctac ggaagcccta tctcgtcatg ggtgtcgagg accgcatcgg caggacctct     660
250 atccctatcg accccgagaa ggttgtaggc atcctcgaat ccgactacca ggacgccact     720
252 gccccaaacg ccgaggccga cgagagtgcg aacaagattg ctggccactt gattgagttc     780
254 ttcgagcacg aggtcgccca cggccgtctc ccgaactccc tccttcccct ccagtccggc     840
256 atcggaacg tcgccaacgc catcatcggg ggctctgaca actccaactt ccgcaacctc     900
258 aagggtctgga ctgaggttat ccaggacacc ttctctgacc tcttcgactc gggccgcctc     960
260 gactttgcca cggccacctc tatccgcttc tcccccgacg gtttccgccg gttctacgac     1020
262 aactgggagg cctactacgg caagctcctc ctccgcagcc agcagggtgc caactcgccc     1080
264 gagatcatcc gccgccttgg tgtcattgcc atgaacaccc ccgtcgagggt cgacatctac     1140
266 gcccacgcca actccacctg cgtcatgggc tcgcgcagtgc tcaacggcct gggcggtccc     1200
268 gccgacttcc tgcgtcctc caagtactct atcatgcaca ccccgctccac ccgcccctcc     1260
270 aagaccgacc cgcacggcgt ctcgtgcac gtcccattgt gcaccacat cgaccagact     1320
272 gagcacgacc tcgacgtcat cgtcaccgag cagggcctgg ccgacgtgcg cggcctgagc     1380

```

RAW SEQUENCE LISTING

DATE: 12/08/2005

PATENT APPLICATION: US/10/527,552

TIME: 10:41:00

Input Set : A:\32677_sequence listing.txt

Output Set: N:\CRF4\12082005\J527552.raw

```

274 cccaggggaga gggcccgcg catcatcaag aagtgcgccc acccggtcta ccagcccatc 1440
276 ctgacccact actttgagaa ggccgagagc gactgcctac gcaagggctg gggccacgag 1500
278 ccccatctgc ttttcaactc gtttgacctg cacaaggccc tcgtggagca cggaagcatg 1560
280 cagaaggctc ggcagtgg 1578
283 <210> SEQ ID NO: 4
284 <211> LENGTH: 2570
285 <212> TYPE: DNA
286 <213> ORGANISM: Acremonium chrysogenum
288 <400> SEQUENCE: 4
289 gaagatcgca tttgggcgag tgggctaata atgcctgctg cctgcctgtg gacggtaaata 60
291 gaattagggtg gaattgtgcg caaattgagg ggaatggccc ccttatcata taaagtgcc 120
293 atgcgatact atggcggtgc gtggggtgcg gtcggtgtcc ggccggtcga acggagggtcc 180
295 cggtatcaa taggcggtag gccggcattg aatcggttcc accgtattcc agacacccaa 240
297 ggaaggcccg ccacccccag ctccggcctg gggatagcgc cgagtggagc actcacgggg 300
299 gccgtgtttg actcgaagac gcgtcgtgat tggccagaac ttcaccccc tctgccaagt 360
301 attggttcac gggattcggc gacgtcaacg accccaccgg ccggattac ataaggtgca 420
303 ctgcagctac tacgtagtac tcgtacttgg gaaggaggga cccttggggg cgagggtttt 480
305 aaaggcaatg gcttcttcgc tgggtccacc aacctgactc tctctctccc ttttacctcg 540
307 ctctctgat tattccctcg tctgcgtctg gatttcatct ctttccctc ccggccctt 600
309 tggatctctg ctctccctc ctctctccc cgcattggtg tgtaaaacca ctgtcccgcg 660
311 gcctcgcgac gagtgcgta ctgcaagccg aaacctcaca atcccttct cacaatggca 720
313 tcaccaatag cctctgccgc cctcaaggcg cgcattcgcc gccctctgat gctcaagaag 780
315 ctgtgcaagc cccaggattt gatgcacac tccccaatg gctcgtacat tggctgggtcc 840
317 ggcttcaccg gcgtcggcta cccgaagtga gtccaccgt catcccgctc cacagtaggc 900
319 gcagccggcc cgctgacagt ccccgacagg aaaatgccga cctacatggc cgatcacgtc 960
321 gagcagaacg gccttcaggg caagctgaag tactcgctat tcgtgggccc atcgctgggt 1020
323 gctgagacag agaatcgctg ggcgctgcct gacatgattg ataggaggac ccgcatcag 1080
325 gtcggcaagg ccatctccaa gggcatcaat gagggcaaga tccacttctt cgacaagcat 1140
327 ctctccatgt tccccgtgga ccttgtatac gtacgtcaac gatgatccct tggaatgtgc 1200
329 atgtactacg agtacctggc gctaacatcc ggtcagggtc actacacaaa ggatagacc 1260
331 cacaacaagc tggacgtggt ggtggtggag gccaccgaca tcaaaggaga cggaagcatt 1320
333 gtaccgggag cttcagtcgg cgcgaccccc gagctcatcc agatggcga taaggtgagc 1380
335 aatttcgatt tctagcggag ggcgcagcag gacctgacat ctccctgtgc agatcattat 1440
337 cgagggtcaac acctcactgc cttcattcga gggctccac gacatcacca tgaccgacct 1500
339 gcccccgcta cggaagccct atctcgcat ggggtgtcag gaccgcatcg gcaggacctc 1560
341 tatccctatc gaccccgaga aggttgtagg catcctcgaa tccgactacc aggacgccac 1620
343 tgcccccaac gccgaggccg acgagagtgc gaacaagatt gctggccact tgattgagtt 1680
345 cttcgagcac gaggtcgccc acggcgtct cccgaactcc ctcttcccc tccagtccgg 1740
347 catcggaac gtgcgaacg ccatcatcgg tggcctcgac aactccaact tccgaacct 1800
349 caaggctctg actgagggtta tccaggacac ctctctcgac ctcttcgact cgggcccgtc 1860
351 cgactttgcc acggccaact ctatccgctt ctcccccgac ggtttccgcc ggttctacga 1920
353 caactgggag gcctactacg gcaagctcct cctccgcagc cagcagggtg ccaactcgcc 1980
355 cgagatcatc cgccgccttg gtgtcattgc catgaacacc ccgctcgagg tcgacatcta 2040
357 cgccacgcc aactccacct gcgtcatggg ctgcgcgatg ctcaacggcc tgggcggtc 2100
359 cgccgacttc ctgcgtcct ccaagtactc tatcatgcac acccgtcca cccgcccctc 2160
361 caagaccgac ccgcacggcg tctcgtgcat cgttcccatg tgcacccaca tcgaccgagc 2220
363 tgagcacgac ctgcagtcga tcgtcaccca gcagggctg gccgacgtgc gcggcctgag 2280
365 cccaggggag agggcccgcg tcatcatcaa gaagtgcgcc caccgggtc accagcccat 2340
367 cctgaccac tactttgaga aggcgagag cgactgccta cgcaagggtc ggggccacga 2400

```

RAW SEQUENCE LISTING

DATE: 12/08/2005

PATENT APPLICATION: US/10/527,552

TIME: 10:41:00

Input Set : A:\32677_sequence_listing.txt

Output Set: N:\CRF4\12082005\J527552.raw

```

369 gccccatctg ctcttcaact cgtttgacct gcacaaggcc ctcgaggagc acggaagcat 2460
371 gcagaagggtc gggcagtggt aagattggcg agacgggaga ggcgttggtg taggagttgg 2520
373 aactagaatc agatatacac cctttcatat atgtagataa tggagccatt 2570
376 <210> SEQ ID NO: 5
377 <211> LENGTH: 16032
378 <212> TYPE: DNA
379 <213> ORGANISM: Acremonium chrysogenum
382 <220> FEATURE:
383 <221> NAME/KEY: misc_feature
384 <222> LOCATION: (15909)..(15909)
385 <223> OTHER INFORMATION: n is a, c, g, or t
387 <400> SEQUENCE: 5
388 tacgtacata cgtcgcgggg ggtagacaat ggtgtggtgt acgtgtacaa ctacagtcag 60
390 acatggacgc aggaacgcga tcattgatac atgcacacgg ggcagcaaat ttagcctggt 120
392 tcactacatg tacatagagg gtacactcca gagcatactg atgggagaaa aagggttcga 180
394 ttgctggttg tttaacatag ccggcaaggg gaaaaaaa agggggcgga gaaggactga 240
396 ttcttcctgg cagacactcg acccttccgg ccccttgaac tgcttttact cccgcattcc 300
398 tccgcacgcc cgcccacagc ggcagatcag ccgaacctga tcgaccgatt taggtctgac 360
400 cattcttggt gatcaagccc cgcagtcctc cctgcagata ctctccgtag gagatggccg 420
402 gcttgctcct ggcgccatcc ttggcccat ccttggcggg cgcgggggtc cacggctgga 480
404 tgggtgctct ccagcccagg ttgacgaaga agggcagtg ctggcgctcc tcgttgaccc 540
406 atttgacgcg gtggatcggg gccgggtagt agtcgtcggg gatatgggcc atgtagctgc 600
408 cgcagttgat gaggaagccc gtgtcgtcag cctggatgtc ctgccagccc tgcggggtct 660
410 tgacctgcag attctgcacg tcggactggg acaacaccgt gatgagggac acgtcctcgt 720
412 gccactcgaa gctgagcttg gtgccgtcgt cggccgtctt gatggccggc tccgggtacg 780
414 ggtcgaggta cgggtaacgg atgagcacga ccgacgagag cgtcgtgtca cggcgggagt 840
416 ggcgggtgaa gaagtcctcg tcgcgacctt gggcgagagc gtagccgcgc agcaccgcgg 900
418 aggagaggcc gaagacgtcc cagtagtact tctcggcgaa ggcccggaac cccgggtgct 960
420 tcgcctcgtc cggccagacg ttgacctcgt gcataggggt gggtccttg attcgcgggt 1020
422 ggtctggggt gaaggagggt ttcaggtagc agaacgattc gaccgccttc ttgcccggga 1080
424 tcggcaggta gtgcccgc cggatctggg actcgtgtc cttgttgtag gcccggtagg 1140
426 cgagctgcca cttctcctcg tccgtgatgc tcatgtggaa ttgttctgc tcgcgcgaga 1200
428 gccacggcag gtgcgacacc tgggttcacc cgtaaaagaa gcctgtgtcg cgcgagtcgg 1260
430 cgtcgtatgg gcgagctacc tcgagcttct tctccttgtc atcgccgaat aggggcgaga 1320
432 catcgattcg ggggacgttg gccactggaa ctggaacgga acccatggtg acggtttgtc 1380
434 ctgcctggtg taagatgtga aagacgagat atgcgtgagt gacgatggcg gaaggagaag 1440
436 cctcgaaaat cagaagagcg accaaaggga tattcaagta ttgcacctc ttgaagctgt 1500
438 ttatacgggc ggctgggtgt gtgtatgtgt acttgagtac ctacctcgtg tctcccgttg 1560
440 ctatacgata tgagcttccc caccagcgcg ctttatggcc tgaccaaggt ctcgattatc 1620
442 cggctcctgc gggtgacact gccgaggggg gttacatacg gtccagcagc ggcgatggag 1680
444 tttgggtcct gaagactgca tggcgggggc aagcgatgag gaacgccgtt acatgcatgt 1740
446 gcatgtagac gccgccaccc atgaggcccg gaacagtcta tcgaagctca gggattggcc 1800
448 cggcaactcg acgcccgcgt gagcggtcca ccggtagtcg acggcgctcg tcggaatctc 1860
450 gcgctgctgc gggccaccac ggcgatgggc cgtacacact gctactacgg tgtacaatgt 1920
452 atcatgtacc cgatcgacga ggaactcggg gtagaggtac cccgtacaat ccagtttctc 1980
454 aacccaatgg aaccacacat acggggtggc ttggttcac gttgcacttt aaactcgcag 2040
456 acgaggggac gacctgcagc gtggccact tctgaagcct gccagcttt ctgcaagacg 2100
458 cggggccatc cgcttggccg aggagagaaa gggatccat ggcgacaaag gcggtcctgg 2160
460 tgggttcggt gccggctttg gagttcactg gtctgggtgg gtggccagct ggatgcatgc 2220

```

RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 12/08/2005

PATENT APPLICATION: US/10/527,552

TIME: 10:41:01

Input Set : A:\32677_sequence listing.txt

Output Set: N:\CRF4\12082005\J527552.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 15909

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:8,9,10,11,12,13,14,15,16,17,18,19,20,21

VERIFICATION SUMMARY

DATE: 12/08/2005

PATENT APPLICATION: US/10/527,552

TIME: 10:41:01

Input Set : A:\32677_sequence_listing.txt

Output Set: N:\CRF4\12082005\J527552.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:918 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:15900